

# Reduced form models of bond portfolios

Matti Koivu\*      Teemu Pennanen†

November 16, 2010

## Abstract

We derive simple return models for several classes of bond portfolios. With only one or two risk factors our models are able to explain most of the return variations in portfolios of fixed rate government bonds, inflation linked government bonds and investment grade corporate bonds. The underlying risk factors have natural interpretations which make the models well suited for risk management and portfolio design.

## 1 Introduction

Bond portfolios are subject to a large number of risk factors. Even in the case of default-free fixed rate bonds, the portfolio return depends in general on the whole yield curve which is a high-dimensional object. Portfolio managers often describe returns in terms of *yield to maturity*, which provides a *one-dimensional* approximation of the expected return on a given portfolio. Such simple descriptions are easy to understand and to model, which makes them useful in risk management and asset allocation.

This paper derives low-dimensional return formulas for several classes of bond portfolios, including fixed rate government bonds, index linked bonds and defaultable corporate bonds. The underlying risk factors in our models have natural interpretations which facilitates the assessment and modeling of the risks. For example, the returns on inflation linked bonds are quite accurately described in terms of the yield and the underlying consumer price index. Returns on corporate bonds, in turn, are well approximated by the yield and the yield spread between corporate and government bonds. With only one- or two risk factors, our models are able to describe consistently over 97% of monthly return variations on government, inflation linked and corporate bonds over the past decades including the recent financial crisis.

Our models are based on low-order Taylor-approximations of the logarithmic price of the portfolio. This is analogous to Chance and Jordan [1996] where a Taylor-approximation of the price with respect to yield and time was found to

---

\*Finnish Financial Supervisory Authority, Market and Operational Risk Division, [matti.koivu@bof.fi](mailto:matti.koivu@bof.fi). The views presented in this paper are those of the authors and are not necessarily shared by the Finnish Financial Supervisory Authority

†Department of Mathematics and Systems Analysis, Aalto University, [teemu.pennanen@tkk.fi](mailto:teemu.pennanen@tkk.fi), corresponding author

give good approximations of return in the case of fixed-rate government bonds. Our model goes one step further by developing Taylor-approximations with respect to time, the yield as well as the outstanding coupon/principal payments. Variations in the outstanding payments are an essential return component e.g. in the case of index linked or corporate bonds. For inflation linked bonds, it amounts upto 10% increase in the explained historical return variations. Even in the case of default free fixed-rate bonds, our models improve on earlier ones. Compared to return formulas obtained by low-order approximations of the price itself, the logarithmic approach simplifies the formulas and improves the accuracy of the approximations.

## 2 Returns on bond portfolios

Consider a bond portfolio whose coupon and principal payments are, as they are received, reinvested in bonds at current market prices. The portfolio's *total return index*  $P_t$  at time  $t$  then equals the market value of the fund's assets at all times. The *yield to maturity*,  $Y_t$  of the portfolio is defined through the equation

$$P_t = \sum_{n=1}^N e^{-Y_t(t_n-t)} C_{t,n}, \quad (1)$$

where  $C_{t,n}$  denotes the fund's outstanding aggregate coupon and principal payments payable at time  $t_n > t$ . Whereas the default-free yield curve can be used to price an arbitrary portfolio of default free bonds, the yield to maturity is defined for a particular portfolio of (default-free or not) bonds. The yield to maturity of a portfolio is sometimes called the "internal rate of return". The number  $N$  of payment dates  $t_n$  in (1) can be allowed to be infinite but in that case our subsequent analysis requires that the outstanding payments  $C_{t,n}$  be bounded in  $n$ . An infinite  $N$  may be useful when modeling perpetual bonds.

The definition of the outstanding payments is crucial in the definition of the yield. In the case of fixed rate bonds, there is little ambiguity: the outstanding payments are what the issuer has promised to deliver even though in the presence of default risk, the actual amount that will be received at time  $t_n$  may be less than  $C_{t,n}$ . Default risk is reflected as a lower price and thus a higher yield. In the case of index linked bonds, the coupon payments are uncertain by definition and the outstanding payments in (1) need to be forecast somehow. The simplest option is to "freeze" the underlying index to its current value when estimating  $C_{t,n}$  at time  $t$ . In the case of inflation linked bonds, the corresponding yield is known as the *real yield*. For now, it suffices to assume that the yield  $Y_t$  is defined according to (1) whatever the definition of outstanding payments may be.

Equation (1) expresses the bond price as a function of time  $t$ , the yield  $Y_t$  and the sequence  $C_t = (C_{t,n})_{n=1}^N$  of payments outstanding at time  $t$ . Indeed, we have  $P_t = P(t, Y_t, C_t)$ , where  $P$  is the function defined for each  $t$ ,  $Y$  and  $C = (C_n)_{n=1}^N$  by

$$P(t, Y, C) = \sum_{n=1}^N e^{-Y(t_n-t)} C_n.$$

Similarly, the log-return on the portfolio over a holding period  $[t, s]$  can be

expressed as

$$\Delta \ln P \approx \ln P(s, Y_s, C_s) - \ln P(t, Y_t, C_t).$$

The approximation is exact if there are no payments during the holding period, i.e. if  $t_1 \geq s$ . If  $s > t_n$ , the  $n$ th payment has been collected and reinvested in the portfolio. An error may result if the reinvested payment appreciates during  $[t_n, s]$  at a rate different from  $Y_s$ . If the payments during  $[t, s]$  amount to a small fraction of the all outstanding payments (if length of the holding period is small compared to the maturity of the bonds in the portfolio), the resulting error could be expected to be small.

Our reduced form models are based on Taylor-approximations of the log-return with respect to time, the yield and the outstanding payments. The first order approximation can be written as

$$\Delta \ln P \approx \frac{1}{P_t} \frac{\partial P}{\partial t}(t, Y_t, C_t) \Delta t + \frac{1}{P_t} \frac{\partial P}{\partial Y}(t, Y_t, C_t) \Delta Y + \sum_{n=1}^N \frac{1}{P_t} \frac{\partial P}{\partial C_n}(t, Y_t, C_t) \Delta C_n \quad (2)$$

where  $\Delta Y = Y_s - Y_t$  and  $\Delta C_n = C_{s,n} - C_{t,n}$ . The first two terms are quite familiar. Indeed, we have

$$\frac{1}{P_t} \frac{\partial P}{\partial t}(t, Y_t, C_t) = Y_t \quad \text{and} \quad \frac{1}{P_t} \frac{\partial P}{\partial Y}(t, Y_t, C_t) = -D_t,$$

where

$$D_t = \frac{1}{P_t} \sum_{n=1}^N (t_n - t) e^{-Y_t(t_n - t)} C_{t,n}$$

is the *Macaulay duration* of the portfolio at time  $t$ . The first term of (2) represents the *pull to par* effect, which means that bonds tend to appreciate when the maturity is approached. The second term gives the price sensitivity with respect to the yield. The effects of time and yield changes on returns off fixed-rate government bonds have been studied e.g. in Chance and Jordan [1996] who developed a Taylor series approximation for the price itself (as opposed to its logarithm).

The last term in (2) is the return component arising from changes in outstanding payments during the holding period  $[s, t]$ . Because of its multivariate character, the last term is difficult to model in general. However, when there are no portfolio updates during the holding period, there is often a single risk factor  $K_s$  such that

$$\Delta C_n \approx K_s C_{t,n} \quad (3)$$

for all  $n = 1, \dots, N$ . For default-free fixed coupon bonds,  $\Delta C_{t,n} = 0$  so (3) holds trivially with  $K_s = 0$ . In the case of index linked bonds, on the other hand,  $K_s$  equals the change in the underlying index over the holding period  $[t, s]$ ; see Section 4 below. When (3) holds, the last term in (2) reduces to

$$\sum_{n=1}^N \frac{1}{P_t} \frac{\partial P}{\partial C_n}(t, Y_t, C_t) \Delta C_{t,n} \approx K_s \frac{1}{P_t} \sum_{n=1}^N e^{-Y_t(t_n - t)} C_{t,n} = K_s$$

and (2) becomes

$$\Delta \ln P \approx Y_t \Delta t - D_t \Delta Y + K_s. \quad (4)$$

The portfolio return over the holding period  $[t, s]$  is then approximated by

$$\frac{\Delta P}{P_t} \approx \exp(Y_t \Delta t - D_t \Delta Y_t + K_s) - 1,$$

where the right side is bounded from below by  $-1$  just like actual portfolio return we are trying to approximate. This is why we linearized the logarithm of the price instead of the price itself as is often done.

The logarithmic approach is supported also by the second order analysis. Denoting the convexity by

$$C_t = \frac{1}{P} \sum_{n=1}^N (t_n - t)^2 e^{-Y(t_n - t)} C_{t,n},$$

we get, by straightforward differentiation, that

$$\frac{\partial^2 \ln P}{\partial t \partial Y} = 1, \quad \frac{\partial^2 \ln P}{\partial Y^2} = C_t - D_t^2, \quad \frac{\partial^2 \ln P}{\partial Y \partial C_n} = \frac{e^{-Y(t_n - t)}}{P} (D_t - t_n + t)$$

while all other second order derivatives of the logarithmic price are zero. If (3) holds, the second order approximation of the log-return reduces (after simple algebraic manipulations) to

$$\Delta \ln P \approx Y_s \Delta t - D_t \Delta Y + K_s + \frac{1}{2} (C_t - D_t^2) \Delta Y^2 \quad (5)$$

This differs from the first order formula (4) only by the addition of the term quadratic in  $\Delta Y$  and in that, in the first term,  $Y_t$  has been replaced by  $Y_s$ . This should be compared with the more complicated return formula of Chance and Jordan [1996] who studied fixed-rate bonds through a second order approximation of the price itself (instead of its logarithm). The reduction of the second order terms in the quadratic approximation of the log return suggests that the logarithm of the price is well approximated already by the first order terms.

### 3 Fixed rate government bonds

In the case of fixed rate default-free bonds, we have  $\Delta C_{t,n} = 0$  so that (3) holds with  $K_s = 0$  and (5) reduces to

$$\Delta \ln P \approx Y_s \Delta t - D_t \Delta Y + \frac{1}{2} (C_t - D_t^2) \Delta Y^2. \quad (6)$$

This is similar to the model studied e.g. in Ilmanen [1992] but there the time component was ignored. Chance and Jordan [1996] incorporated the time component but their model, like that of Ilmanen [1992], was based on a Taylor-approximation of the price instead of its logarithm.

#### 3.1 Empirical results

We study the accuracy of the above models in explaining monthly returns on fixed rate government bonds. Our dataset covers end of month observations of the Barclays' market capitalisation weighted total return indices, durations and yields for France, Germany, Italy, United Kingdom, United States and the Euro

area<sup>1</sup>. The length of the time series for each country is given at the bottom of Table 1.

We fit the following two models to the data

$$\begin{aligned}\text{Model 1: } \Delta \ln P_s &= c + Y_s \Delta t - D_t \Delta Y_s + \epsilon_s, \\ \text{Model 2: } \Delta \ln P_s &= c + Y_s \Delta t - D_t \Delta Y_s + \gamma \Delta Y_t^2 + \epsilon_s.\end{aligned}$$

The parameters  $c$  and  $\gamma$  are estimated by ordinary least squares. The regression statistics provide us with diagnostic tools to evaluate the performance of the proposed models. By comparing the fits of the two models we can assess the significance of the quadratic term in explaining the returns.

Table 1 displays the estimation results for the two model specifications. Model 1 already provides an almost perfect fit to the total return data with  $R^2$  values ranging from 99.6% to 99.8%. The model fit has been consistent over time and the residuals have remained marginal even during the recent financial crisis. The  $R^2$  values being close to 100% there is not much room for improvement when adding the second order term. The estimated coefficients  $\gamma$  in Model 2 are generally significant at a 5% level but, consistently with previous research [Chance and Jordan, 1996], the improvements in the  $R^2$  statistics are marginal, less than 0.02% points in all the studied markets. The Partial- $R^2$  statistic is defined as the  $R^2$ -statistic obtained by regressing the residual of Model 1 with the quadratic term. The quadratic term of Model 2 explains less than 5.11% of the residual variance of Model 1. The estimated constant  $c$  in Model 1 deviates from zero at 5% significance level, but with varying signs. This may indicate that the model has not entirely captured all the systematic return components. The addition of the quadratic term in Model 2 mitigates this effect to some extent.

Table 1: Estimation results for fixed rate government bonds.

	FRA	GER	IT	UK	US	EURO
<b>Model 1</b>						
100 * $c$	0.0147 (3.5842)	0.0083 (1.7215)	0.0084 (1.9339)	-0.0195 (-4.8403)	-0.0202 (-3.0658)	0.015 (4.1467)
$R^2$	99.76%	99.64%	99.68%	99.84%	99.71%	99.79%
<b>Model 2</b>						
100 * $c$	0.007 (1.4255)	0.0007 (0.1186)	-0.0001 (-0.0178)	-0.0256 (-5.7509)	-0.0277 (-3.6689)	0.01 (2.2687)
$\gamma$	24.4889 (2.7949)	24.1545 (2.2625)	32.5155 (2.5933)	4.8396 (3.0699)	11.1262 (1.9703)	17.031 (1.9131)
$R^2$	99.77%	99.65%	99.70%	99.85%	99.72%	99.79%
Partial- $R^2$	5.11%	3.41%	4.43%	2.63%	2.84%	2.46%
Data start	1997-12	1997-12	1997-12	1980-12	1998-12	1997-12
Data end	2010-03	2010-03	2010-03	2010-03	2010-03	2010-03

Notes. The table contains estimation results for the models:

Model 1:  $\Delta \ln P_s = c + Y_s \Delta t - D_t \Delta Y_s + \epsilon_s$ ,

Model 2:  $\Delta \ln P_s = c + Y_s \Delta t - D_t \Delta Y_s + \gamma \Delta Y_t^2 + \epsilon_s$ ,

Numbers in parentheses are the  $t$ -statistics for the estimated coefficients.

<sup>1</sup>Further information is available online at <https://ecommerce.barcap.com/indices/index.dxml>

Our results are consistent with Ilmanen [1992] who finds that the explanatory power of duration has increased over time. In his empirical study, Ilmanen [1992] found that the duration explained 80% to 90% of return variance of fixed rate government bonds during the 1980s. The nearly 100%  $R^2$ -values for Model 1 during the past decade suggest that the trend has continued. It should be noted, however, that our results are not strictly comparable since our model explains the log-returns and it includes the time component.

## 4 Inflation linked bonds

The coupon and principal payments of most inflation linked bonds are tied to an underlying consumer price index so that the payment that will be received at time  $t_n$  is  $C_{t_n,n} = I_{t_n}C_{0,n}$ , where  $C_{0,n}$  is the *real payment* at time  $t = 0$  and  $I_{t_n}$  is the value of the underlying consumer price index some time before  $t_n$ . The specification of the *indexation lag* is specified in the contract and depends e.g. on the lag between time for which the index is computed and the time when its value is reported. We refer the reader to [Deacon et al., 2004, Chapter 2] for a general overview of different cash flow structures on index linked bonds. The *real yield* of a portfolio of inflation linked bonds is defined by setting the outstanding payments according to the most recent value of the consumer price index  $I_t$ , i.e.

$$C_{t,n} = I_t C_{0,n}$$

in (1). It follows that

$$\begin{aligned} \Delta C_n &= I_s C_{0,n} - I_t C_{0,n} \\ &= \left( \frac{I_s}{I_t} - 1 \right) C_{t,n} \\ &= \pi_s \Delta t C_{t,n}, \end{aligned}$$

where  $\pi_s$  is the *annualized rate of inflation* over the period  $[t - \delta, s - \delta]$  where  $\delta$  denotes the indexation lag. Assumption (3) thus holds with  $K_s = \pi_s \Delta t$ , so the first-order approximation (4) can be written as

$$\Delta \ln P \approx (Y_s + \pi_s) \Delta t - D_t \Delta Y. \quad (7)$$

### 4.1 Empirical results

In our empirical study, we consider the two model specifications

$$\begin{aligned} \text{Model 1: } \Delta \ln P_s &= c + Y_s \Delta t - D_t \Delta Y_s + \epsilon_s, \\ \text{Model 2: } \Delta \ln P_s &= c + (Y_s + \pi_s) \Delta t - D_t \Delta Y_s + \epsilon_s, \end{aligned}$$

the first one of which ignores the third term in the Taylor-approximation (2). This allows us to evaluate the significance of the changes in the outstanding payments (the  $K$ -term in (4)) in explaining the portfolio returns.

We use monthly observations of Barclays inflation linked government bond index data. The data consists of total return indices, yields and durations on portfolios of inflation linked government bonds for Canada, France, South Africa, Sweden, United Kingdom and United States<sup>2</sup>. The bonds' cash flows are

---

<sup>2</sup>Further information is available online at <https://ecommerce.barcap.com/indices/index.dxml>

linked to the evolution of country specific inflation indices (usually the general consumer price index). The monthly time series of appropriate inflation indices were obtained from Eurostat and the national statistical authorities' websites. Since the indexation lag is not specified in the data, we select the lag  $\delta$  that gives the best fit to historical data. In all the studied countries, the data strongly supports a specific choice of the lag. The length of the time series for each country is given at the bottom of Table 2.

Table 2 gives the regression statistics for the two model specifications in the six markets. The inclusion of the  $K$ -term in Model 2 results in a substantial improvement over Model 1. Model 2 gives an extremely good fit to the total return data with  $R^2$  values ranging from 97.6% to 99.8%. The lowest  $R^2$  value for South Africa is mainly caused by one outlier observation caused by pricing distortions in the local conventional bond market in 2002 and the subsequent inversion of the breakeven inflation curve.

The high values of Partial- $R^2$  statistics mean that the  $K$ -term is able to explain most of the residual variance of Model 1. Figure 1 depicts the unexplained residual returns for the two models. The reduction in the residuals between Model 2 and Model 1 is caused solely by the incorporation of the inflation term.

The fact that the estimated constant terms in Model 2 are generally insignificant is another indication that the model is well specified. The country specific "optimal" lag lengths  $\delta$  are given on the last row of Table 2.

Table 2: Regression statistics for inflation linked government bonds.

	CAN	FRA	SA	SWE	UK	US
<b>Model 1</b>						
100*c	0.1642	0.1328	0.4935	0.1114	0.2073	0.1966
t-stat	(6.201)	(5.544)	(9.367)	(3.678)	(7.611)	(5.888)
$R^2$	97.48%	95.46%	78.6%	91.29%	96.06%	94.41%
<b>Model 2</b>						
100*c	0.0046	-0.0098	-0.0069	0.0087	-0.0151	-0.0004
t-stat	(0.628)	(-0.839)	(-0.389)	(1.465)	(-1.918)	(-0.066)
$R^2$	99.80%	98.91%	97.58%	99.66%	99.67%	99.82%
Partial- $R^2$	92.22%	76.02%	88.69%	96.14%	91.43%	96.78%
Lag	2	2	3	2	1	3
Data start	1996-12	1998-09	2000-03	1996-12	1996-01	1997-03
Data end	2010-03	2010-03	2010-03	2010-03	2010-03	2010-03

Notes. The table contains estimation results for the models:

Model 1:  $\Delta \ln P_s = c + Y_s \Delta t - D_t \Delta Y_s + \epsilon_s$ ,

Model 2:  $\Delta \ln P_s = c + (Y_s + \pi_{s-k}) \Delta t - D_t \Delta Y_s + \epsilon_s$ ,

Numbers in parentheses are the  $t$ -statistics for the estimated coefficients.

## 5 Corporate bonds

It was found in Ilmanen et al. [1994] that duration's explanatory power in explaining bond returns quickly decreases when moving from default-free bonds to corporate bonds. This section derives a simple two-factor model for the returns of well-diversified portfolios of non-callable corporate bonds. Our aim is to show

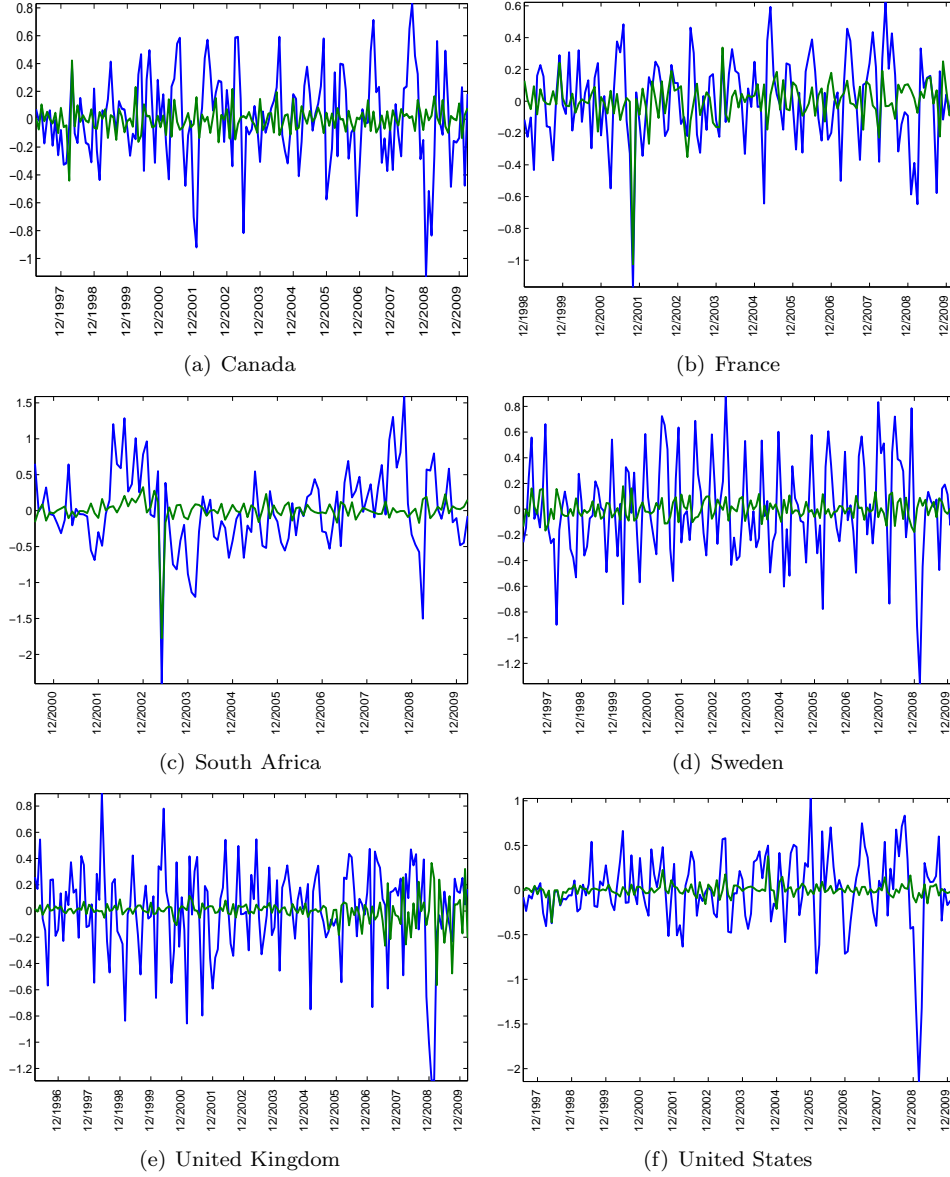


Figure 1: Residual returns (in percentages) of Model 1 (blue) and Model 2 (green) for inflation linked bonds.

that the effect of defaults can be approximated by (3) where  $K_s$  can, in turn, be approximated by the yield spread between corporate and government bonds. We start by a mathematical justification based on familiar assumptions on the default process and on the diversification of the portfolio. The model will then



be validated by an empirical analysis of Merrill-Lynch data on corporate bond portfolios.

Because of defaults, the outstanding payments  $C_{t,n}$  of a portfolio of corporate bonds may change during the holding period  $[t, s]$ . We will use the *multiple defaults* model described e.g. in Schönbucher [1998] and [Schönbucher, 2003, Chapter 6]. The multiple defaults assumption can be justified by restructuring of defaulting issuers that often takes place in practice. Thus, we assume that each default event reduces the outstanding payments  $C_{t,n}$  to a fraction  $\rho C_{t,n}$  of their pre-default values. The number  $\rho$  is called the *recovery rate*, which in the context of portfolio modeling represents the reduction in the *portfolio's* outstanding payments due to an individual default. A crucial assumption in justifying (3), is that each default event reduces the outstanding payments  $C_{t,n}$  for all  $n = 1, \dots, N$  by, approximately, the same recovery rate.

The time of occurrence and the recovery rate associated with the  $i$ th default will be denoted by  $t_i$  and  $\rho_i$ , respectively. The sequence  $(t_i, \rho_i)_{i=1}^\infty$  then determines the outstanding payments remaining after default events during a time interval  $[t, s]$ . Indeed, we have

$$C_{s,n} = \prod_{i=1}^{I_s} \rho_i C_{t,n},$$

where  $I_s = \max\{i \mid t_i \in [t, s]\}$  is the number of default events during  $[t, s]$ . Both the default times  $t_i$  and the recovery rates  $\rho_i$  are random in general. In order to allow dependencies between them, we will model  $(t_i, \rho_i)_{i=1}^\infty$  as a Cox process in  $[0, \infty) \times [0, 1]$  with intensity  $\lambda_{t,\rho}$ ; see e.g. Schönbucher [1998] or Gaspar and Slinko [2008]. That is, we assume that, conditionally on  $\lambda$ , the number of default events with  $(t_i, \rho_i)$  in a measurable subset  $B$  of  $[0, \infty) \times [0, 1]$  is Poisson distributed with parameter

$$\int_B \lambda_{t,\rho} dt d\rho$$

and the number of events in disjoint subsets of  $[0, \infty) \times [0, 1]$  are independent. One could model defaults by a general *point process* (see e.g. [Kallenberg, 2002, Chapter 12]), but for notational simplicity, we assume that the process can be expressed in terms of an intensity. Note that the default times follow a Cox process with intensity  $\lambda_t = \int_0^1 \lambda_{t,\rho} d\rho$ . The above model thus extends the one studied e.g. in Lando [1998]; see Schönbucher [1998] for further discussion and references.

It follows (see the Appendix) that the amount of outstanding payments due at time  $s \geq t$  has conditional expectation

$$E[C_{s,n} \mid \lambda] = C_{t,n} \exp \left( - \int_t^s l_u du \right), \quad (8)$$

where

$$l_u = \int_0^1 (1 - \rho) \lambda_{u,\rho} d\rho$$

is the *mean loss rate*. Note that we can write  $l_t = \lambda_t(1 - \rho_t)$ , where  $\lambda_t = \int_0^1 \lambda_{t,\rho} d\rho$  and  $\rho_t = \int_0^1 \rho \lambda_{t,\rho} d\rho / \lambda_t$  for  $\lambda_t > 0$  and  $\rho_t = 0$  otherwise. This corresponds to the defaultable term structure models proposed e.g. in Duffie and Singleton [1999].

Arguing like in Jarrow et al. [2005] (who studied defaultable zero coupon bonds with zero recovery), one can show that the default event risk can be diversified away in the sense that the outstanding payments  $C_{s,n}$  converge to  $E[C_{s,n} | \lambda]$  when the number of issuers in the portfolio is increased; see the Appendix. This suggests that for a well-diversified portfolio, the effect of default events on outstanding payments can be approximated by

$$C_{s,n} \approx C_{t,n} \exp \left( - \int_t^s l_u du \right) \quad (9)$$

and thus that (3) holds with

$$K_s = \exp \left( - \int_t^s l_u du \right) - 1.$$

In order to estimate  $K_s$ , we will assume that there exists a *risk neutral measure*  $Q$  under which the market prices of traded securities are equal to the expectations of their discounted cash flows; see e.g. [Duffie and Singleton, 2003, Chapter 5]. The default-free short rate will be denoted by  $r$ . The amount of cash that will be received at time  $t_n$  is equal to  $C_{t_n,n}$ , the outstanding payments at time  $t_n$  payable at time  $t_n$ . Applying (9) with  $s = t_n$ , we get

$$\begin{aligned} P_t &= E^Q \sum_{n=1}^N \exp \left( - \int_t^{t_n} r_u du \right) C_{t_n,n} \\ &\approx E^Q \sum_{n=1}^N \exp \left( - \int_t^{t_n} r_u du \right) \exp \left( - \int_t^{t_n} l_u du \right) C_{t,n} \\ &= E^Q \sum_{n=1}^N \exp \left( - \int_t^{t_n} r_u^c du \right) C_{t,n}, \end{aligned}$$

where  $r_t^c = r_t + l_t$  is the *defaultable short rate*. This is analogous to the reduced form pricing formulas obtained e.g. in Schönbucher [1998], Duffie and Singleton [1999]. The instantaneous loss rate can thus be expressed as the *short spread*  $l_t = r_t^c - r_t$ .

Unfortunately, neither  $r_t^c$  nor  $r_t$  is observable in practice. We will approximate them by yields on short maturity corporate and government bond portfolios, respectively. This suggests the approximation

$$K_s \approx \exp(-S_s \Delta t) - 1, \quad (10)$$

where  $S_s$  is the short spread at time  $s$ . When  $S_s \Delta$  is small, we can simplify this to

$$K_s \approx -S_s \Delta t. \quad (11)$$

Figures 2 and 3 display historical monthly values of  $K_s$  and  $S_s \Delta t$  in percentages for European and US corporate bond markets during 1996/2–2010/1. The values of  $K_s$  were obtained by solving  $K_s$  from (4) for observed values of log-returns and yields on Merrill-Lynch investment grade corporate bond indices. In the figures, the spread  $S_s$  is the difference between yields on portfolios of corporate and government bonds of 1-3 years to maturity. The values of  $S_s \Delta t$  have remained below 0.7% which means that the relative error in using (11) instead of (10) would have been less than 0.0025% points.

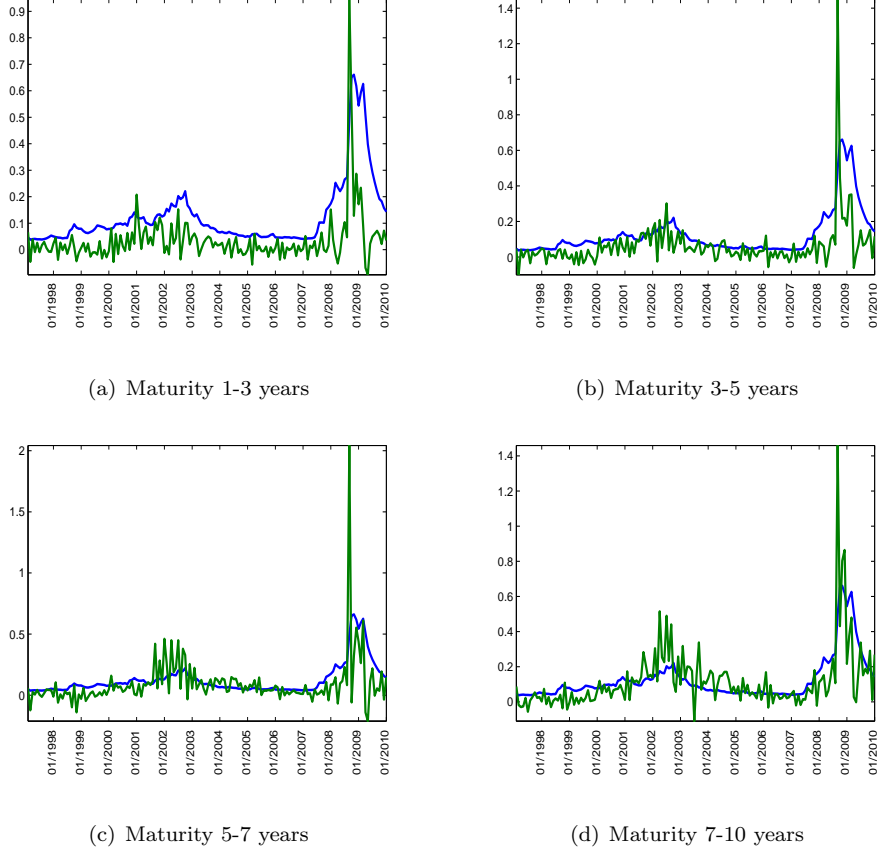


Figure 2: Historical evolution (in percentages) of  $S_t \Delta t$  (blue) and residual returns of Model 1 representing default losses (green) during 1996/2–2010/1 for the US indices.

Using (11) in (4) suggests the approximation

$$\Delta \ln P \approx (Y_s - \alpha S_s) \Delta t - D_t \Delta Y \quad (12)$$

for log-returns on corporate bond portfolios. We allow the constant  $\alpha$  to deviate from one since even with the shortest maturity available (1-3 years in our numerical examples below), the spread  $S_s$  is likely to overestimate the mean loss rate  $l_s$ . Indeed, the yield on corporate bonds contains a premium for unexpected variations in the future loss rate. In addition to this *systematic default risk* there may be a premium for *default event risk* which remains when the diversification argument used to derive (9) is not exactly satisfied; see the Appendix.

## 5.1 Empirical results

We will again consider two different model specifications

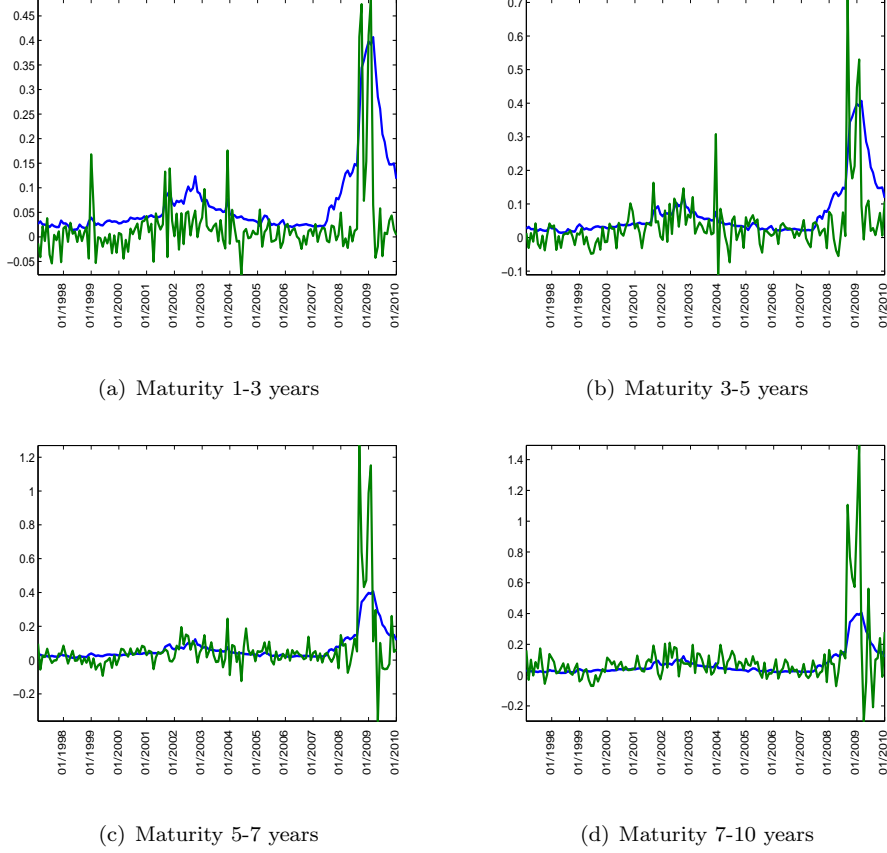


Figure 3: Historical evolution of  $S_t\Delta t$  (blue) and residual returns of Model 1 representing default losses (green) during 1996/2–2010/1 for the European indices.

$$\begin{aligned} \text{Model 1: } \Delta \ln P_s &= c + Y_s\Delta t - D_t\Delta Y_s + \epsilon_s, \\ \text{Model 2: } \Delta \ln P_s &= c + (Y_s - \alpha S_s)\Delta t - D_t\Delta Y_s + \epsilon_s, \end{aligned}$$

which differ by the term  $\alpha S_s$ , where  $S_s$  is the yield spread on short maturity bond indices. The parameter  $\alpha$  will be estimated from the data. We fit the models to monthly observations of Merrill-Lynch investment grade corporate bond indices data. Our dataset covers the total return indices, yields and durations of market capitalization weighted investment grade corporate bond portfolios of maturities 1–3, 3–5, 5–7 and 7–10 years from the US and European markets. As a proxy for the instantaneous loss rate we will use the yield spread of 1–3 year short maturity bonds.

Table 3 displays the estimation results for the European and Table 4 for the US data under the two model specifications. Model 1, which ignores the effects of default losses on the outstanding payments, already explains the majority of the return variations with  $R^2$  values ranging between 96.4% and 99.2%. The

Table 3: Return regressions for the EU corporate bond portfolios.

	Model 1				Model 2			
	1-3 Y	3-5 Y	5-7 Y	7-10 Y	1-3 Y	3-5 Y	5-7 Y	7-10 Y
$100 * c$	-0.022 (-3.431)	-0.038 (-5.014)	-0.064 (-4.357)	-0.094 (-5.754)	0.013 (1.811)	0.007 (0.818)	0.028 (1.753)	0.012 (0.675)
$\alpha$					0.509 (7.424)	0.633 (8.246)	1.312 (8.921)	1.502 (9.426)
$R^2$	96.38%	98.63%	97.5%	98.21%	97.33%	99.05%	98.35%	98.87%
Partial- $R^2$					26.35%	30.63%	34.07%	36.58%
Data start	1997-1	1997-1	1997-1	1997-1	1997-1	1997-1	1997-1	1997-1
Data end	2010-1	2010-1	2010-1	2010-1	2010-1	2010-1	2010-1	2010-1

*Notes.* The table contains estimation results for two different model specifications:

Model 1:  $\Delta \ln P_s = c + Y_s \Delta t - D_t \Delta Y_s + \epsilon_s$ ,

Model 2:  $\Delta \ln P = (Y_s - \alpha S_s) \Delta t - D_t \Delta Y_s + \epsilon_s$ ,

Numbers in parentheses are the  $t$ -statistics for the estimated coefficients.

Table 4: Return regressions for the US corporate bond portfolios.

	Model 1				Model 2			
	1-3 Y	3-5 Y	5-7 Y	7-10 Y	1-3 Y	3-5 Y	5-7 Y	7-10 Y
$100 * c$	-0.033 (-3.506)	-0.061 (-5.265)	-0.09 (-5.424)	-0.116 (-8.145)	0.02 (1.754)	0.011 (0.747)	-0.001 (-0.035)	-0.001 (-0.098)
$\alpha$					0.423 (6.462)	0.571 (7.169)	0.712 (6.005)	0.915 (10.643)
$R^2$	97.81%	98.51%	98.35%	99.19%	98.27%	98.88%	98.66%	99.53%
Partial- $R^2$					21.33%	25.02%	18.97%	42.38%
Data start	1997-1	1997-1	1997-1	1997-1	1997-1	1997-1	1997-1	1997-1
Data end	2010-1	2010-1	2010-1	2010-1	2010-1	2010-1	2010-1	2010-1

*Notes.* The table contains estimation results for two different model specifications:

Model 1:  $\Delta \ln P_s = c + Y_s \Delta t - D_t \Delta Y_s + \epsilon_s$ ,

Model 2:  $\Delta \ln P = (Y_s - \alpha S_s) \Delta t - D_t \Delta Y_s + \epsilon_s$ ,

Numbers in parentheses are the  $t$ -statistics for the estimated coefficients.

addition of the the spread term in Model 2 inmpoves the model fit in all maturities and markets. The spread term is able to explain roughly 20–40% of the residual return variation of Model 1. In addition to improving the models' fit the addition of the spread terms also dilutes the significance of the estimated constant terms, which may be interpreted as a sign of an improved model specification. These quantitative results are well supported by Figures 2 and 3 that display the evolution of unexplained residual returns of Model 1 together with the short maturity yield spread in the US and European markets, respectively.

The estimates of  $\alpha$  as well as the Partial- $R^2$  values tend to increase as a function of the index maturity. This can be explained, to a large extent, by the differences in the rating composition and seniority of the bonds included in the indices. Figure 4 displays the evolution of the rating distributions of the

US corporate bond indices by maturity bucket during the worst phases of the financial crises, when the default losses reached exceptionally high levels. The rating quality deteriorates with maturity which explains the increased default losses at the longer maturities. The composition of the US indices by instrument type are displayed in Figure 5. In line with the deterioration of the rating quality, the seniority of the bonds contained in the indices also declines as the maturity of the indices increases. The largest default losses were observed in the 5–7 year index which had the largest share of capital related debt instruments (under the Basel II regulation) and securitized exposures, both of which exhibited large losses during the financial crisis. On the other hand, the shorter end of the maturity spectrum contained the highest share of senior debt and thus exhibited the lowest default losses. Similar conclusions regarding the deterioration of the rating quality of the indices as a function of the maturity hold for the European indices as well but we omit the details.

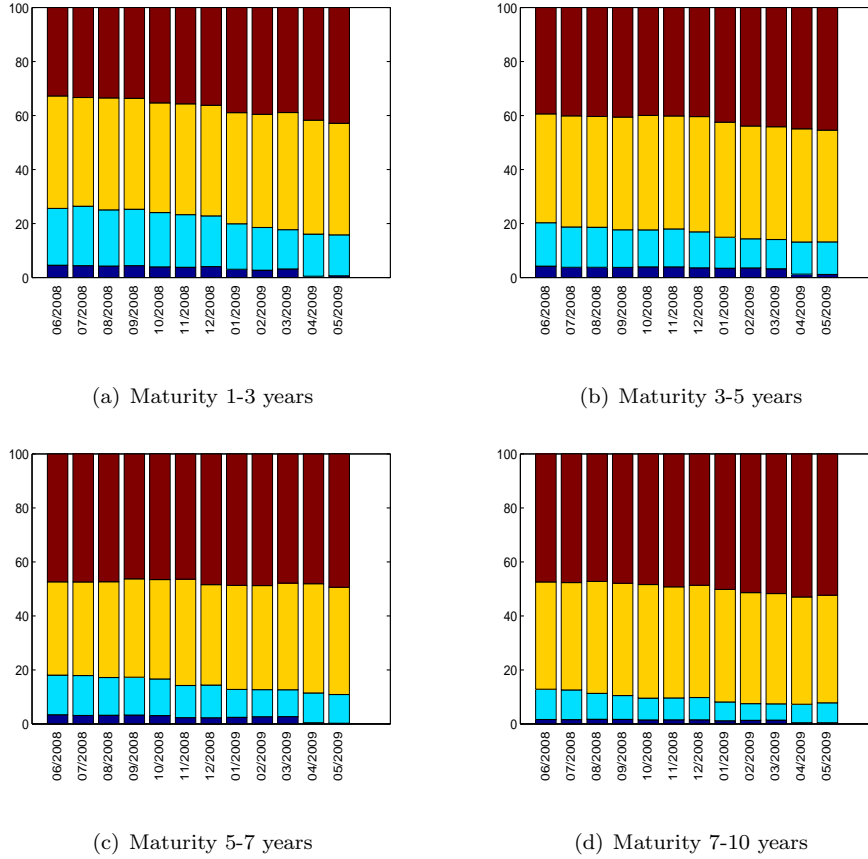
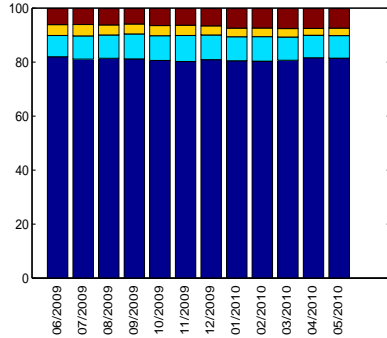
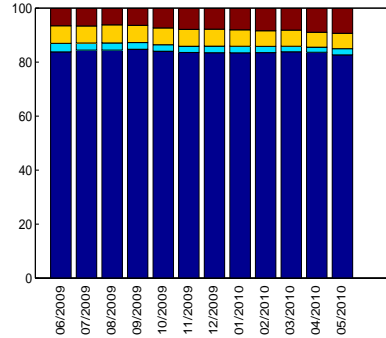


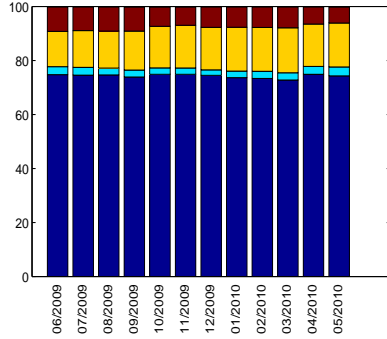
Figure 4: The composition of US corporate bond indices by rating classes AAA (dark blue), AA (light blue), A (yellow) and BBB (red) during 6/2009-5/2010.



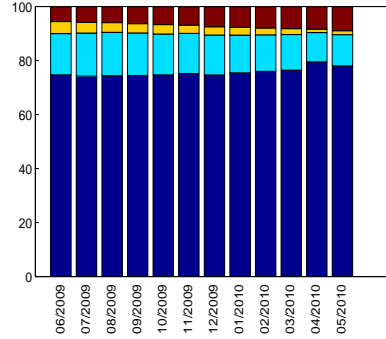
(a) Maturity 1-3 years



(b) Maturity 3-5 years



(c) Maturity 5-7 years



(d) Maturity 7-10 years

Figure 5: The distribution of US corporate bond indices among senior debt (dark blue), subordinated debt (light blue), capital instruments (yellow) and securitized debt (red) during 6/2009-5/2010.

The three terms in Model 2 explain majority of the return variation in the considered eight portfolios but, especially at the longer maturities, there remains residual spikes in the aftermath of the Lehman Brothers' collapse in September 2008. Some idiosyncratic default event risk thus seems to remain, even though the indices contain bonds from hundreds of issuers. The systematic default risk on the other hand seems to be captured to a large extent by the spread term.

## Appendix

We prove (8) by using the notion of *Laplace functional* of a random measure; see e.g. [Kallenberg, 2002, Chapter 12]. To this end, we write

$$\begin{aligned} C_{s,n} &= \prod_{i=1}^{I_s} \rho_i C_{t,n} \\ &= \exp \left( \sum_{i=1}^{I_s} \ln \rho_i \right) C_{t,n} \\ &= \exp \left( - \int_t^\infty \int_0^1 f d\mu(u, \rho) \right) C_{t,n}, \end{aligned}$$

where  $f(u, \rho) = -\chi_{[t,s]}(u) \ln \rho$  and  $\mu = \sum_i \delta_{t_i, \rho_i}$ . Here  $\chi_{[t,s]}$  denotes the characteristic function of the interval  $[s, t]$  and  $\delta_{t_i, \rho_i}$  is the Dirac measure that assigns unit mass to the point  $(t_i, \rho_i)$ . Conditionally on  $\lambda$ , the process  $(t_i, \rho_i)_{i=1}^\infty$  is Poisson (since it is a Cox process), so by [Kallenberg, 2002, Lemma 12.2(i)],

$$\begin{aligned} E[C_{s,n} | \lambda] &= \exp \left( - \int_t^\infty \int_0^1 (1 - e^{-f(u, \rho)}) \lambda_{u, \rho} du d\rho \right) C_{t,n} \\ &= \exp \left( - \int_t^s \int_0^1 (1 - \rho) \lambda_{u, \rho} du d\rho \right) C_{t,n} \\ &= \exp \left( - \int_t^s l_u du \right) C_{t,n}, \end{aligned}$$

where

$$l_u = \int_0^1 (1 - \rho) \lambda_{u, \rho} d\rho.$$

This gives (8).

The approximation (9) can be justified by assuming as in Jarrow et al. [2005], that the portfolio is diversified among a large number  $M$  of issuers so that

$$C_{t,n} = \sum_{i=1}^M w^i C_{t,n}^i,$$

where  $w^i$  is the relative weight and  $C_{t,n}^i$  is the outstanding payments of the  $i$ th issuer, respectively. The defaults of each issuer are assumed to follow a Cox process with intensity  $\lambda$  as described in Section 5. If, conditionally on  $\lambda$ , the defaults of different issuers are independent, we get

$$\begin{aligned} E(C_{s,n} - E[C_{s,n} | \lambda])^2 &= EE[(C_{s,n} - E[C_{s,n} | \lambda])^2 | \lambda] \\ &= E \sum_{i=1}^M (w^i)^2 E[(C_{s,n}^i - E[C_{s,n}^i | \lambda])^2 | \lambda] \\ &\leq E \sum_{i=1}^M (w^i C_{t,n}^i)^2, \end{aligned}$$

where the inequality holds since both  $C_{s,n}^i$  and  $E[C_{s,n}^i | \lambda]$  are bounded between zero and  $C_{t,n}^i$ . Assuming as in Jarrow et al. [2005], that  $C_{t,n}^i$  are uniformly bounded in  $i$  and that the portfolio is well-diversified in the sense that  $\sum_{i=1}^M (w^i)^2 \rightarrow 0$  as  $M$  grows, we get that  $C_{s,n}$  converges in mean square to  $E[C_{s,n} | \lambda]$ .



## References

- D.M. Chance and J.V. Jordan. Duration, convexity, and time as components of bond returns. *The Journal of Fixed Income*, 6(2):88–96, 1996.
- M. Deacon, A. Derry, and D. Mirfendereski. *Inflation-indexed Securities: Bonds, Swaps and Other Derivatives*. John Wiley & Sons, 2004.
- D. Duffie and K.J. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720, 1999.
- D. Duffie and K.J. Singleton. *Credit Risk: Pricing, Measurement and Management*. Princeton University Press, 2003.
- R. M. Gaspar and I. Slinko. On recovery and intensity correlation – a new class of credit risk models. *The Journal of Credit Risk*, 4(2):1–33, 2008.
- A. Ilmanen. How well does duration measure interest rate risk? *The Journal of Fixed Income*, 1(4):43–51, 1992.
- A. Ilmanen, D. McGuire, and A. Warga. The value of duration as a risk measure for corporate debt. *The Journal of Fixed Income*, 4(1):70–76, 1994.
- R. A. Jarrow, D. Lando, and F. Yu. Default risk and diversification: Theory and empirical implications. *Mathematical Finance*, 15(1):1–26, 2005.
- O. Kallenberg. *Foundations of Modern Probability*. Springer-Verlag, New York, 2002.
- D. Lando. On cox processes and credit-risky securities. *Review of derivatives research*, 2:99–120, 1998.
- P. J. Schönbucher. Term structure modelling of defaultable bonds. *Review of Derivatives Research*, 2(2–3):161–192, 1998.
- P. J. Schönbucher. *Credit Derivatives Pricing Models: Model, Pricing and Implementation*. John Wiley & Sons Ltd, West Sussex, 2003.